

The Fresh-Finger Property ^{*}

John Howat¹, John Iacono², and Pat Morin³

¹ School of Computing
Queen's University
`howat@cs.queensu.ca`

² Department of Computer Science and Engineering
Polytechnic Institute of New York University
`jiacono@poly.edu`

³ School of Computer Science
Carleton University
`morin@scs.carleton.ca`

Abstract. The *unified property* roughly states that searching for an element is fast when the current access is close to a recent access. Here, *close* refers to rank distance measured among all elements stored by the dictionary. We show that distance need not be measured this way: in fact, it is only necessary to consider a small working-set of elements to measure this rank distance. This results in a data structure with access time that is an improvement upon those offered by the unified property for many query sequences.

1 Bounds for searching: notation and history

Comparison-based searching is one of the most fundamental operations in computer science: given a set S of n totally ordered items, create a data structure that, given a query key x , will return the largest key in S that is no larger than x . This is called *predecessor search*. We focus on the case where S is static, and thus can be assumed to be the integers from 1 to n . We refer to a search that returns x as an *access* to x . Let $A = \langle a_1, a_2, \dots, a_m \rangle$ denote a sequence of accesses to be performed on a data structure, with m chosen to be sufficiently large to absorb any start-up costs.

Any comparison-based search data structure is, at its core, a method of choosing which comparisons to perform in order to execute an access. The data structure is essentially a way of encoding a comparison tree to execute each access—the data structure could do this in an explicit way like in a binary search tree, or in an implicit way as in the binary search algorithm. It has been long-known that information theory tells us that the worst-case time for an access must be $\Omega(\log n)$, and that $O(\log n)$ can be achieved with data structures such as binary search on a sorted array.

But, worst-case analysis is not the end of the story, as one can design data structures that execute operations in $o(\log n)$ time if the operations have some

^{*} This research was partially supported by NSERC and MRI.

kind of order to them. Thus, we can create data structures with access times that are functions of the access sequences themselves (or some distributional statistic of the access sequence)—these runtimes will still have $O(\log n)$ worst-case behavior, but will be faster over sequences that have certain desirable characteristics. We now review some runtime bounds that have been introduced, and the data structures whose runtimes have bounds (we will say a data structure has a bound to mean that its runtime can be bounded by the bound):

Static optimality bound. If the number of searches in A to x is $f(x)$, then the runtime to search for a_i is $O\left(\log \frac{m}{f(a_i)}\right)$.⁴ Knuth showed how to achieve this bound if the $f(\cdot)$ -values are given in advance [1].

Working-set bound. Let $w_i(x)$ be the number of distinct items accessed since the last access to x in a_1, \dots, a_{i-1} . A data structure has the working-set bound if an access to a_i takes time $O(\log w_i(a_i))$. The idea behind this is that if the accesses are restricted to a subset of k items, then the accesses will take time $O(\log k)$ rather than $O(\log n)$. It has been shown that the working-set bound implies the static optimality bound in the amortized sense. Splay trees [2] have the working-set bound in the amortized sense, while the working-set structure [3] was designed to have this bound in the worst-case.

Queueish bound [4]. The working-set bound requires that items which were accessed recently take less time than those that have not been accessed in a while. The queueish bound reverses this and states that the time to access a_i should be $O(\log(n - w_i(a_i)))$; thus any structure with the queueish bound will execute the least recently accessed item in constant time. No dictionary is known to have the queueish bound and it remains open whether such a dictionary can exist; however, it was shown that there is a structure with a close-to-queueish bound of $O(\log n - w_i(a_i) + \log \log n)$ amortized access time.

Dynamic finger bound [5,6]. Let $d(x, y)$ be the number of keys between x and y in S (this is just $|x - y|$ if S is the integers from 1 to n). The dynamic finger property says the cost to execute search a_i is $O(\log d(a_{i-1}, a_i))$. Level-linked trees [7] have the dynamic finger property in the worst-case, and splay trees have the amortized dynamic finger property.

Unified bound [8]. The dynamic finger bound and the working-set bounds are the best known bounds on the runtime of splay trees, yet neither implies the other and neither is tight. For example in the sequence $\langle 1, 2, \dots, n, 1, 2, \dots \rangle$ the dynamic finger will give a bound of $O(1)$ per operation on average, while the working-set bound gives a bound of $O(\log n)$ per operation. In the sequence $\langle 1, n, 1, n, \dots \rangle$, the situation is reversed. The unified bound was proposed as a

⁴ In this paper, $\log 0 = \log 1 = 1$.

natural combination of these two bounds;⁵ an access is fast if it is close in key value to something that has been recently accessed. Formally, a data structure has the unified bound if accessing a_i takes time $O(\log \min_j (d(a_i, j) + w_i(j)))$. This clearly implies the working-set bound (set $j = a_i$) and the dynamic finger bound (set $j = a_{i-1}$). A non-tree structure was presented with the unified bound, and it is conjectured that splay trees have the unified bound. A binary search tree (BST) structure with the unified bound plus an additive $O(\log \log n)$ is known [9], and a BST structure with the unified bound was claimed [10], but later declared to be buggy [11].

There are several issues that are important when considering a bound:

Static vs. dynamic. If a search algorithm uses the same search tree for every access, it is said to be *static*, while if the comparisons performed to execute a given search depend upon the previous searches performed it is said to be *dynamic*. The static optimality bound is the best bound possible if the search algorithm generates the same comparison tree for every access.

Online. A bound where the runtime bound to execute a_i is a function of the sequence $\langle a_1 \dots a_i \rangle$ is said to be an online bound. All of the bounds listed above, except the static optimality, bound are online. The static optimality bound is not online because it is computed as a function of the frequency count over the entire length of a sequence.

Amortization. For any operation there can be at most $O(2^k)$ different searches than can be done using at most k comparisons. Any bound that at any time that has $\omega(2^k)$ different searches perform only k comparisons for some value of k means that the bound can not hold in the worst case. None of the bounds above require amortization, and if a bound does require amortization, that is probably a sign that it is somehow unnatural.

Binary Search Tree model. Wilber [12] formalized the binary search tree model; in this model the data structure is a binary search tree which can be restructured through the use of rotations. The set of sequences which binary search trees can execute quickly seems to be a reasonable classifier of those sequences that we consider to be natural. Without a restriction to the BST model, given any single access sequence, it is possible to create a data structure that will execute the searches in that sequence quickly, and others slowly. This is not possible in the BST model as there are deterministic sequences such as the bit reversal permutation that can not be executed faster than $O(\log n)$ amortized time.

The class of BST data structures also have the possibility that there may exist an online BST data structure that can execute every access sequence asymptotically as fast as the best BST data structure *for that sequence*. Such a structure

⁵ In an unfortunate naming conflict, Sleator and Tarjan have a “Unified Theorem” for splay trees [2, Theorem 5] and the bound in the Unified Theorem is also sometimes called the “unified bound.”

would be called *dynamically optimal*; no BSTs are known to be dynamically optimal although spay trees and Lucas’ trees [13] are conjectured to be. Blum et. al. [14] gave a non-tree data structure that runs within a constant factor of comparisons of any BST data structure, but requires superpolynomial time to decide which comparisons to perform. Tango trees [15] are a BST that execute every sequence within a $O(\log \log n)$ factor of the best possible binary search tree. Of the bounds described above, no BST can have the queueish property, it is conjectured that there is a BST with the unified property, and the rest of the bounds described above are achievable by BST data structures.

2 Problems with the unified bound

The unified bound is the best proposed bound for binary search trees, and seems to be a reasonable combination of temporal locality and locality in keyspace. However, we will show that the unified bound has a flawed view of keyspace, and propose a new bound that attempts to rectify this flaw.

Recall that the unified property roughly states than access is fast when the current access is close to a recent access. For example, consider the following access sequence (assume n is even and n divides m):

$$A = \left\langle 1, \frac{n}{2} + 1, 2, \frac{n}{2} + 2, 3, \frac{n}{2} + 3, \dots, \frac{n}{2}, n \right\rangle^{m/n} \quad (1)$$

(The exponentiation denotes that the sequence is repeated m/n times to make a sequence of length m .) Observe that, except for the first two accesses in each cycle, every access is at distance one from the element accessed two accesses ago. A dictionary with the unified property would therefore perform this sequences in time at most

$$2(m/n) \cdot O(\log n) + (n - 2)(m/n) \cdot O(1) \in O(m)$$

for an amortized cost of $O(1)$ per access.

Next, consider the following access sequence:

$$A' = \left\langle K, \frac{n}{2} + K, 2K, \frac{n}{2} + 2K, 3K, \frac{n}{2} + 3K, \dots, \frac{n}{2}, n \right\rangle^{mK/2n} \quad (2)$$

where n is a multiple of K , mK is a multiple of n , and $n^{1/4} \leq K \leq \sqrt{n}$. For this sequence, the unified bound is useless: any element accessed less than $K/2$ time units in the past is at distance at least K from the currently accessed element, so the cost of every access is $\Omega(\log K) = \Omega(\log n)$.

On the other hand, the sequence A' is not very different from A . Indeed A' can be viewed as the sequence A over a larger set, S , in which a $(1 - 1/K)$ fraction of the elements are never accessed. Intuitively, in a good data structure these irrelevant elements should “fall out of the way” in order to speed up accesses to the important elements (multiples of K).

The sequence A' demonstrates the problem with the unified property: The distance function $d(x, y)$ simply measures the number of keys between x and y . But, suppose some key values have not been accessed in a long time relative to x and y , or in the extreme case, have never been accessed. Why should the number of such keys between x and y influence the runtime of accessing keys such as x and y ? Put simply, they should not. Data structures such as splay trees will have items that are never accessed “percolate” to the bottom of the structure, and the runtime of a splay tree with a subtree of never-accessed keys is identical to the runtime if the keys are not there. Thus we need a more nuanced $d(\cdot, \cdot)$ function that “forgets” keys that have not been accessed in a while when computing key distance.

In this paper, we will expand on the idea of counting only recently accessed elements towards the distance between elements stored by the dictionary. The remainder of the paper is organized in the following way. In Section 3, we define a new, stronger version of the unified property. In Section 4, we show how to construct a dictionary that has this new property. We conclude with possible directions for future research in Section 5.

3 Defining the fresh-finger property

In this section, we define a new, stronger version of the unified property that we term the *fresh-finger property*. Recall that $A = \langle a_1, a_2, \dots, a_m \rangle$ is our access sequence. Define

$$l_i(x) = \min(\{\infty\} \cup \{j > 0 \mid a_{i-j} = x\})$$

One can think of $l_i(x)$ as the most recent time x has been queried in A before time i . We then define

$$w_i(x) = \begin{cases} n & \text{if } l_i(x) = \infty \\ |\{a_{i-l_i(x)+1}, \dots, a_i\}| & \text{otherwise} \end{cases}$$

which is the *working-set number* of x at time i . We also define $W_i(j)$ to be the set of all elements $x \in S$ at time i such that $w_i(x) \leq j$, i.e., the set of all elements with working-set number at most j at time i . Next, we define $d_T(x, y)$ to be the rank distance between a and b in the set T , i.e.,

$$d_T(x, y) = \begin{cases} |\{z \in T : x < z \leq y\}| & \text{if } x < y \\ |\{z \in T : y < z \leq x\}| & \text{otherwise.} \end{cases}$$

Finally, define

$$y_i(x, T) = \arg \min_{y \in T} w_i(y) + d_T(x, y)$$

We are now ready to define the fresh-finger property. In terms of the preceding notation, the unified property states that the time to access the element $x \in S$ at time i is

$$O(\log(w_i(y_i(x, S)) + d_S(x, y_i(x, S)))) \quad (3)$$

A first attempt at defining the fresh-finger property might be

$$O(\log(w_i(y_i(x, W_i(w_i(x)))) + d_{W_i(w_i(x))}(x, y_i(x, W_i(w_i(x)))))) . \quad (4)$$

Equation (4) should be contrasted with the definition of the unified property defined by (3). In (4), the rank distance between x and other elements is measured only with respect to the set $W_i(w_i(x))$, the set of elements that have been accessed since the last access to x (i.e., a set of *fresh fingers*). In (3), the rank distance is measured with respect to S , the entire set of elements stored in the data structure.

Since $W_i(w_i(x)) \subseteq S$, (4) is certainly a stronger requirement. Unfortunately, it is too strong, and there is no comparison-based data structure that can achieve this bound in the worst-case. To see this, consider the access sequence

$$\langle 1, 2, 3, \dots, n, 1, x \rangle$$

where $x \in \{1, \dots, n\}$. Let $i = n + 2$ (so that $a_i = x$ represents the second access to x). Then $W_i(w_i(x)) = \{1, x, x+1, x+2, \dots, n\}$. But then the rank difference, $d_{W_i(w_i(x))}(x, 1)$, between x and 1 is at most 1 and $w_i(1) = 1$, so, according to (4), the time to accessing x is at most

$$O(\log(w_i(1) + d_{W_i(w_i(x))}(x, 1))) = O(1)$$

But this is true for any $x \in \{1, \dots, n\}$, so for *any* of the n choices for x , (4) requires that a data structure execute the access in constant time. This is not information-theoretically possible; accessing a randomly chosen $x \in \{1, \dots, n\}$ requires at least $\log_2 n$ comparisons in expectation.

From the preceding discussion, we conclude that the set in which we measure rank distance should be expanded. This leads to the following definition:

Definition 1. *A data structure has the fresh-finger property if its runtime for an access is bounded by*

$$O(\log(w_i(y_i(x, W_i(w_i(x)^2)))) + d_{W_i(w_i(x)^2)}(x, y_i(x, W_i(w_i(x)^2)))) .$$

Observe that the set $W_i(w_i(x)^2)$ contains elements that have been accessed less recently than x . These additional elements will allow us to support the rest of the access cost while respecting information-theoretic lower bounds. The intuition for expanding the set under consideration in this manner is the fact that the data structure will consist of substructures that increase doubly-exponentially in size, and so by squaring the working-set number under consideration, we take advantage of elements in an adjacent substructure.

For brevity, we define

$$y_i(x) = y_i(x, W_i(w_i(x)^2))$$

and

$$\text{FF}_i(x) = \log(w_i(y_i(x)) + d_{W_i(w_i(x)^2)}(x, y_i(x)))$$

As an example, consider the following access sequence, where 15 is the element currently being accessed at the end of the sequence.

$$2, 3, 4, 5, 6, 7, 8, 9, \overbrace{10, 11, 12, 13, 14, 15, 16, 1}^{W_i(w_i(15)^2)}, 15$$

$W_i(w_i(15))$

The original definition of the fresh-finger property uses $W_i(w_i(15))$, while the modified definition uses $W_i(w_i(15)^2)$. This modified definition allows for 9, 10, 11, 12, 13 to contribute to the rank distance. This results in a query time that does not violate information-theoretic lower bounds, since it does not result in a situation where all n queries must be executed in constant time. Before presenting a data structure that (nearly) achieves the fresh-finger property, it is worth doing a sanity-check of this definition. In particular, we confirm that there exists (distributions over) sequences $A = \langle a_1, \dots, a_m \rangle$ such that

$$\text{FF}(a_1, \dots, a_m) = \sum_{i=1}^m \text{FF}_i(a_i)$$

is a lower-bound for accessing a_1, \dots, a_m .

Theorem 1. *For all positive integers $n, r \leq n$, and $m \geq 2r \log n$, there exists a distribution, \mathcal{A} , over $\{1, \dots, n\}^m$ such that, for any comparison-based dictionary data structure, D , that stores $\{1, \dots, n\}$, and an access sequence $A = a_1, \dots, a_m$ drawn from \mathcal{A}*

1. $\text{FF}(a_1, \dots, a_m) = O(m \log r)$.
2. *the expected number of comparisons performed by D while accessing A is $\Omega(m \log r)$.*

Proof. The sequence A is defined as $a_i = i$ for $i \leq r$ or a_i is selected uniformly at random from the set $\{1, \dots, r\}$ for $r < i \leq m$. This choice of A immediately implies that

$$(m - r) \log_2 r \geq (m/2) \log_2 r = \Omega(m \log r)$$

is a lower-bound on the expected number of comparisons performed by D while accessing the (randomly chosen values) a_{r+1}, \dots, a_m . This establishes Part 2 of the result.

On the the other hand, to establish Part 1, we have

$$w_i(a_i) \leq \begin{cases} n & \text{for } i \leq r \\ r & \text{for } r < i \leq m \end{cases}$$

Thus,

$$\text{FF}(a_1, \dots, a_m) \leq r \log n + (m - r) \log(2r) = O(m \log r)$$

since $m \geq 2r \log n$.

4 Towards the fresh-finger property

In this section, we describe a data structure that comes to within a small additive term of achieving the fresh-finger property.

4.1 The data structure

The data structure consists of k finger search trees T_1, T_2, \dots, T_k as well as k accompanying queues Q_1, Q_2, \dots, Q_k . Recall that finger search trees can support insertions and deletions in $O(1)$ worst-case time (when provided with a pointer to the element to be deleted) and finger searches in $O(\log d)$ worst-case time, where d is the distance between the element being searched for and the supplied pointer into the data structure [16].

The size of T_j is 2^{2^j} , except for T_k which has size n . It follows that k is $O(\log \log n)$. We will maintain the invariant that $T_j \subset T_{j+1}$ for all $1 \leq j < n$. The queue Q_j contains exactly the same elements as T_j in the order they were inserted into T_j . Pointers are maintained between elements in the queue and corresponding elements in the finger search tree.

To perform a search, we will perform finger searches for x in T_1, T_2, \dots until we find x for the first time (say, $x \in T_j$). In T_1 , we use an arbitrary element as the starting finger for the search. In all other trees, we run two finger searches for x in parallel: one from the successor of the element found in the previous finger search tree, and one from the predecessor of the element found in the previous finger search tree. As soon as the first of these two searches terminates, we stop the other.

To restructure the data structure after we have found $x \in T_j$, we must insert x into T_1, T_2, \dots, T_{j-1} (note that x is not present in any of these trees, since if it were, it would have already been found) and enqueue x in Q_1, Q_2, \dots, Q_{j-1} . At this point, we note that each of T_1, T_2, \dots, T_{j-1} and Q_1, Q_2, \dots, Q_{j-1} are too big. We therefore dequeue the oldest element in each of Q_1, Q_2, \dots, Q_{j-1} and delete the corresponding elements in T_1, T_2, \dots, T_{j-1} .

4.2 Analysis

Recall that we are aiming for a running time of

$$O(\text{FF}_i(x)) = O(\log(w_i(y_i(x)) + d_{W_i(w_i(x)^2)}(x, y_i(x))))$$

Consider a search for x at time i , and consider the element $y_i(x)$. Suppose x first appears in T_j and $y_i(x)$ first appears in $T_{j'}$. Because x first appears in T_j , we have that $w_i(x) \geq 2^{2^{j-1}}$. Therefore, $j \leq \log \log w_i(x) + O(1)$. Similar reasoning shows $w_i(y_i(x)) \geq 2^{2^{j'-1}}$, so that $j' \leq \log \log w_i(y_i(x)) + O(1)$. We consider three cases, based on how j compares with j' :

If $j \leq j'$ (i.e., x appears no later than $y_i(x)$),⁶ then the running time follows easily: x has working-set number $w_i(x) \leq w_i(y_i(x))$. The element x can thus be found in time $\sum_{l=1}^j 2^l = O(2^j)$, which is $O(2^{j'}) = O(\log w_i(y_i(x)))$.

The more interesting case occurs when $j > j'$ (i.e., x appears after $y_i(x)$). In this case, the algorithm will reach the tree, $T_{j'}$, containing $y_i(x)$ in time

$$\sum_{\ell=1}^{j'} O(\log 2^{2^\ell}) = \sum_{\ell=1}^{j'} 2^\ell = O(2^{j'}) = O(\log w_i(y_i(x))) .$$

The search in $T_{j'}$ finds both the predecessor and and successor, y_1 and y_2 of x in $T_{j'}$. That is,

$$y_1 \leq x \leq y_2$$

and $y_i(x)$ is not in the open interval (y_1, y_2) . In particular, for any set T one of y_1 or y_2 , say y_1 , has

$$d_T(x, y_1) \leq d_T(x, y_i(x))$$

Indeed, from this point onwards, every search in T_ℓ , for each $\ell \in \{j'+1, \dots, j\}$, y'_ℓ such that

$$d_T(x, y'_\ell) \leq d_T(x, y_i(x)) .$$

The elements in $T_{j'+1}, \dots, T_{j-1}$ are all in $W_i(w_i(x))$, and so the remaining searches in $T_{j'+1}, \dots, T_{j-1}$ therefore take a total of at most

$$\begin{aligned} & (j - j' - 1)O(\log d_{W_i(w_i(x))}(x, y_i(x))) \\ &= O((\log d_{W_i(w_i(x))}(x, y_i(x)))(\log \log w_i(x))) \end{aligned}$$

time.

At last, the final search, in T_j is the expensive one, since the only guarantee we have on the elements of T_j are that their working-set number is at most $w_i(x)^2$. Thus, the elements in T_j are a subset of the elements in $W_i(w_i(x)^2)$ and the time to search in T_j is at most

$$d_{W_i(w_i(x)^2)}(x, y_i(x)) .$$

In either case, the total search time thus far is at most

$$\begin{aligned} & O(\log w_i(y_i(x))) \\ &+ O((\log d_{W_i(w_i(x))}(x, y_i(x)))(\log \log w_i(x)) + \log d_{W_i(w_i(x)^2)}(x, y_i(x))) \end{aligned}$$

At this point, x has been found and we must now adjust the data structure. First, x must be inserted in T_1, T_2, \dots, T_{j-1} . Because we have a finger for x inside each of these structures, this takes total time $O(\log \log w_i(x))$. Enqueueing x in each of Q_1, Q_2, \dots, Q_{j-1} also takes $O(j) = O(\log \log w_i(x))$. The subsequent deletions and dequeueings of the oldest elements in Q_1, Q_2, \dots, Q_{j-1}

⁶ In fact, this case can only occur when $j = j'$, since otherwise $w_i(y_i(x)) > w_i(x)$, and so $w_i(y_i(x)) + d(x, y) > w_i(x) + d(x, x)$, which contradicts the definition of $y_i(x)$.

and T_1, T_2, \dots, T_{j-1} take a total of $O(j) = O(\log \log w_i(x))$ time as well, since the dequeuing operation takes $O(1)$ time and provides a pointer to the node in the corresponding tree where the deletion must be performed. Therefore, all restructuring operations take time $O(\log \log w_i(x))$.

We therefore have

Theorem 2. *There exists a static dictionary over the set $\{1, 2, \dots, n\}$ that supports querying element x in worst-case time*

$$O(\mathit{FF}_i(x) + (\log d_{W_i(w_i(x))}(x, y_i(x)))(\log \log w_i(x)))$$

5 Conclusion

In this paper, we defined a stronger version of the unified property and described a data structure that achieves it to within a small additive term. Instead of computing rank distance over the entire dictionary, we compute rank distance only within a working-set containing an element that is close to a recently-accessed element.

There are several possible directions for future research.

1. One can measure distances within the set $W_i(w_i(x)^{1+\epsilon})$ instead of the set $W_i(w_i(x)^2)$ by changing how the substructures grow. Is it possible to reduce this further? For example, is it possible to measure distances within the set $W_i(O(w_i(x)))$?
2. We argued that it is not possible to measure within $W_i(w_i(x))$ in the *worst case*. Is it possible to measure within this set in the amortized sense?
3. Can the additive term in Theorem 2 be reduced? It seems difficult to reduce this term below $\Omega(\log \log w_i(x))$ using an approach similar to the one presented here, since elements must shift through at least this many substructures.
4. We have only considered the case where S is static. Is it possible to maintain the fresh-finger property while supporting insertions into and deletions from S ?

References

1. Knuth, D.E.: Optimum Binary Search Trees. *Acta Inf.* **1** (1971) 14–25
2. Sleator, D.D., Tarjan, R.E.: Self-Adjusting Binary Search Trees. *J. ACM* **32**(3) (1985) 652–686
3. Iacono, J.: Alternatives to splay trees with $O(\log n)$ worst-case access times. In Kosaraju, S.R., ed.: *SODA, ACM/SIAM* (2001) 516–522
4. Iacono, J., Langerman, S.: Queaps. *Algorithmica* **42**(1) (2005) 49–56
5. Cole, R., Mishra, B., Schmidt, J.P., Siegel, A.: On the Dynamic Finger Conjecture for Splay Trees. Part I: Splay Sorting $\log n$ -Block Sequences. *SIAM J. Comput.* **30**(1) (2000) 1–43
6. Cole, R.: On the Dynamic Finger Conjecture for Splay Trees. Part II: The Proof. *SIAM J. Comput.* **30**(1) (2000) 44–85

7. Hoffman, K., Mehlhorn, K., Rosenstiehl, P., Tarjan, R.E.: Sorting Jordan Sequences in Linear Time Using Level-Linked Search Trees. *Information and Control* **68**(1-3) (1986) 170–184
8. Badoiu, M., Cole, R., Demaine, E.D., Iacono, J.: A unified access bound on comparison-based dynamic dictionaries. *Theor. Comput. Sci.* **382**(2) (2007) 86–96
9. Derryberry, J., Sleator, D.D.: Skip-Splay: Toward Achieving the Unified Bound in the BST Model. In Dehne, F.K.H.A., Gavrilova, M.L., Sack, J.R., Tóth, C.D., eds.: *WADS*. Volume 5664 of *Lecture Notes in Computer Science*, Springer (2009) 194–205
10. Derryberry, J.: Adaptive Binary Search Trees. PhD thesis, CMU (2009)
11. Sleator, D.: Achieving the unified bound in the BST model Talk. No proceedings.
12. Wilber, R.E.: Lower Bounds for Accessing Binary Search Trees with Rotations. *SIAM J. Comput.* **18**(1) (1989) 56–67
13. Lucas, J.M.: Canonical forms for competitive binary search tree algorithms. Technical report, Tech. Rep. DCS-TR-250, Rutgers University (1988)
14. Blum, A., Chawla, S., Kalai, A.: Static optimality and dynamic search-optimality in lists and trees. In: *SODA'02: Proceedings of the 13th Annual ACM-SIAM Symposium on Discrete Algorithms*. (2002) 1–8
15. Demaine, E.D., Harmon, D., Iacono, J., Patrascu, M.: Dynamic Optimality - Almost. *SIAM J. Comput.* **37**(1) (2007) 240–251
16. Brodal, G.S., Lagogiannis, G., Makris, C., Tsakalidis, A.K., Tsihlias, K.: Optimal finger search trees in the pointer machine. *J. Comput. Syst. Sci.* **67**(2) (2003) 381–418